Appendix C: Diacritics and Special Characters

BACKGROUND: Diacritics are marks placed over, under, or through letters or combinations of letters that indicate differences in pronunciation and meaning between the marked letter and its unmarked counterpart. Common diacritics are the umlaut (for example, ü), the acute (for example, ó), the grave (for example, à), and the macron (for example, ō). Double diacritics are two diacritics modifying a single letter (for example Hát bội). Special characters occupy a place of their own rather than modifying a letter. Common special characters include the Polish l (ł), the mi͡agkiĭ znak (´), the ayn (ʻ), and the alif (ʼ). Superscript characters are letters, numbers, or symbols written or printed higher on a line of text; subscript characters are letters, numbers, or symbols written or printed lower on a line of text.

PLACEMENT OF DIACRITICS IN LCGFT: In LCGFT diacritics and special characters are generally printed in accordance with either the LC romanization conventions for a particular language or with a conventional form found in reference sources.

1. Policy for established genre/form terms. Transcribe existing terms exactly as they have been established, using diacritics and special characters as indicated.

2. Establishing terms with or without diacritics. When, as a result of doing authority research, a choice must be made between a term without diacritics and the otherwise identical term with diacritics, generally prefer the term without diacritics if it is *clearly* the conventional English form. If there is no conventional English form, prefer the form of the term as found in the *Encyclopaedia Britannica* or in reference sources of the particular discipline or language, using diacritics to conform to LC's transcription or romanization conventions.

3. Diacritic use and LC practice. If it is determined that an authorized term or reference is to have diacritics, use diacritics that conform to LC transcription or romanization conventions, even if reference sources indicate different diacritics. Use judgment in determining whether to make UF references from forms with variant diacritics that are spelled differently from the proposed term. (For forms that have variant diacritics but are spelled the same as the proposed heading, see sec. 4, below.)

4. Diacritics in authorized terms and references. Because some online systems cannot distinguish between a term without diacritics and the otherwise identically spelled term with diacritics, do not add 455 fields for variants of this type, even if the terms do not represent the same word in a particular language.

5. Greek or Hebrew letters. Romanize Greek or Hebrew letters when establishing terms. Do not use Greek or Hebrew letters in authorized terms or references, even though they appear as such in reference sources.

6. Ampersands.   Do not use ampersands in authorized terms; spell out the word **and**. Ampersands may be used in UF references if the form with the ampersand is found in reference sources, in which case a reference with and without spaces around the ampersand should be provided. Provide an additional UF with the ampersand spelled out as the word **and** if that form is not already present in the record. *Example:*

```
FAQs
UF Q & As
UF Q and As
UF Q&As
```

7. Common diacritics and special characters.

a. Acute (´)

```
Guaguancós (Music)
```

> Do not confuse this diacritic with the **mi͡agki͡i znak** (ʹ) and the **apostrophe** ('), which are special characters.

b. Alif (')

c. Angstrom or Circle above letter (°)

> Do not confuse with the **dot above letter** (·).

d. Apostrophe (')

```
Commedia dell'arte
```

> Do not confuse with the **acute** (´), the **mi͡agki͡i znak** (ʹ), the **alif** ('), the **ayn** ('), or the **grave** (`).

e. Ayn (')

> Do not confuse with the **acute** (´), the **mi͡agki͡i znak** (ʹ), the **alif** ('), or the **apostrophe** (')

7. *Common diacritics and special characters*. *(Continued)*

f. Breve (˘)

Do not confuse with the **hacek** (ˇ)

g. Cedilla (¸)

Do not confuse with the **left hook** (˛).

h. Circle below letter (˳)

Do not confuse with the **dot below letter** (̣).

i. Circumflex (^)

j. Dot above letter *or* Superior dot (˙)

Do not confuse with the **angstrom** or **circle above letter** (°).

k. Dot below letter (̣)

```
Vil pāṭṭu
```

Do not confuse with the **circle below letter** (˳).

l. Double acute (″)

This diacritic appears only in Hungarian. Do not confuse with the **tverdyi znak** (″), a special character that follows the letter it modifies, the **Umlaut** (¨), or the standard quotation mark (").

m. Grave (`)

```
Romans à clef
```

Do not confuse with the **apostrophe** (').

7. *Common diacritics and special characters*. *(Continued)*

   n. Hacek (ˇ)

   Do not confuse with the **breve** (˘).

   o. Left hook (˛)

   Do not confuse with the **cedilla** (¸).

   p. Ligature (⁀)

   q. Macron (¯)

   Dādrās

   r. Mi͡agkiĭ znak (′)

   Do not confuse with the **acute** (´), the **alif** (ʼ), the **apostrophe** (ʼ), or the **ayn** (ʻ).

   s. Polish l (ł)

   Note that Polish also has a regular l.

   t. Slash o (ø)

   u. Tilde (~)

   Flamenco music
   UF Cañi, Cante

   v. Turkish i (ı)

   Note that i is *not* dotted. If uppercased, render as I.
   Turkish also has a regular i.

7. *Common diacritics and special characters*. *(Continued)*

w. Tverdyi znak (″)

Do not confuse this special character with the **double acute** (″) of Hungarian, the **umlaut** (¨), or the standard quotation mark (").

x. Umlaut *or* diaeresis (¨)

Fairy tales
UF Märchen

Do not confuse this diacritic with the **double acute** (″) of Hungarian, the **tverdyi znak** (″), or the standard quotation mark (").